THEORETICAL CHEMISTRY INSTITUTE

THE UNIVERSITY OF WISCONSIN

APPLICABILITY OF PERTURBATION THEORY TO MOLECULAR PROBLEMS

by

Joseph O. Hirschfelder

WIS-TCI-135

11 November, 1965

This paper was presented at The Advanced Seminar on "Perturbation Theory and its Application in Quantum Mechanics" sponsored by the University of Wisconsin Mathematics Research Center (U. S. Army) and Theoretical Chemistry Institute on October 4-6, 1965 at Madison, Wisconsin. The proceedings of this symposium, including this paper, will be published by John Wiley and Sons

MADISON, WISCONSIN

APPLICABILITY OF PERTURBATION THEORY TO MOLECULAR PROBLEMS*

By

Joseph O. Hirschfelder

University of Wisconsin Theoretical Chemistry Institute

Madison, Wisconsin

This paper was presented at The Advanced Seminar on "Perturbation Theory and its Application in Quantum Mechanics" sponsored by the University of Wisconsin Mathematics Research Center (U. S. Army) and Theoretical Chemistry Institute on October 4-6, 1965, at Madison, Wisconsin. The proceedings of this symposium, including this paper, will be published by John Wiley and Sons, Inc. (1966).

ABSTRACT

This paper is a general survey of the ways in which low order perturbation theory is used in quantum mechanics to determine the energy and other properties of molecules. The various types of mathematical problems encountered are discussed.

- - - - -

* This research was supported by the National Aeronautics and Space Administration Grant NsG-275-62.

Paper size 4-1/2 x 7-1/4 (incl. r.h.) 27 x 44 picas

APPLICABILITY OF PERTURBATION THEORY TO MOLECULAR PROBLEMS

by

Joseph O. Hirschfelder

University of Wisconsin Theoretical Chemistry Institute

Madison, Wisconsin

Now Fills Thy Sleep With Perturbation (Richard III, Shakespeare)

It is easy to write down a Hamiltonian operator such that its eigenvalues give the energy of each of the allowed states of a molecule and such that the corresponding eigenfunctions determine the probability density of single electrons and clusters of pairs, triples, etc. in electron configuration space. To solve such a Schrödinger equation is comparable in difficulty with the classical problem of determining the precise orbits of each of the planets, moons, comets, etc. in the solar system. Clearly the solution to such problems must be approximated by the use of perturbation series and variational procedures. Theoretical chemists are trying to develop suitable methods of solution. Some of the difficulties which they encounter could be overcome by the use of techniques which are known by the professional mathematicians. It is the purpose of this symposium to open up a communications channel between the theoretical chemists (who have the problems) and the mathematicians (who can diagnose the problems). Actually, many of the new developments in spectral theory are closely related to our practical problems. Thus, we have mutual research interests.

We hope that the development of new mathematical methods and the availability of high-speed computing machines will make it feasible to calculate expectation values of the energy and other properties of molecules with a precision at least comparable to corresponding values obtained by laboratory experimentation. Preferably, we should calculate both upper and lower bounds for each property in order to establish the reliability of our theoretical estimates. However, there are a number of mathematical difficulties which must be overcome before we can succeed in our practical objectives. We want to thank you for coming to this Symposium and helping us to "overcome"!

I will try to outline the ways in which perturbation theory is used in molecular quantum mechanics and point out some of the mathematical problems [1]. Perturbation theory is one of the most promising approaches to molecular quantum mechanics. It has the great advantage that the functional form of the perturbed wave function is shaped by the perturbation itself. Frequently, sufficient accuracy is obtained from even the first-order perturbed wave function. From such a function, the energy can be computed accurate through the third order and the expectation values of other properties can be obtained accurate through the first order. Indeed, the wave function accurate through the n-th order permits the energy to be calculated through the $(2n+1)$-st order.

I. REASONS FOR RECENT SUCCESS

Recent interest in perturbation theory has been sparked by four considerations:

1) A number of methods can be used for the direct solution of the perturbation equations. Formerly, the solutions of the perturbation equations were expressed in a spectral representation as a sum over the discrete energy states and an integral over the continuum energy states of coefficients times eigenfunctions of the unperturbed Hamiltonian. Unfortunately, in most molecular problems, it is difficult to construct a satisfactory unperturbed Hamiltonian for which we know the complete set of eigenfunctions and eigenvalues. Thus, the spectral representation formalism was seldom useful. Instead, we now seek direct solution to the inhomogeneous differential or differential-integral perturbation equations. Sometimes these equations are separable and the solutions can be obtained by quadrature. However, more frequently, the solutions are approximated by the use of variational principles.

2) Good variational principles are available for both upper and lower bounds to particular orders of perturbation energy. These principles are generally applied to appropriate truncated function sets. The resulting optimum linear combinations provide approximations to individual terms in the perturbed wave function. However, the relationship between variational principles and perturbation theory is not limited to this type of application. Every general variational principle applies separately to each individual order of perturbation. Thus, the Rayleigh-Ritz variational principle leads to the Hylleraas principle which gives an upper bound to the second order energy. Similarly, the variational principles associated with the virial, hyper-virial, and Hellmann-Feynman theorems are useful in perturbation theory.

3) The Dalgarno "interchange theorem" can be used for calculating properties other than energy. This greatly simplifies the calculations of the first order correction to the expectation values which result from the "badness" of the approximate wave function. The only perturbation equations which must be solved involve a real or fictitious external perturbation, and not the internal perturbation. Since the external perturbation is usually the sum of one-electron operators, these equations may be easy to solve. Without the Dalgarno "interchange theorem" it would have been necessary to solve much more difficult perturbation equations involving two-electron repulsion terms. The first order correction to the physical properties should result in considerable improvement over the zeroth order expectation values which have previously been available. Unfortunately, the interchange theorem provides very little help in calculating the second order and higher corrections.

4) Wigner showed that if we know the wave functions accurate through the nth order then we can calculate the energy accurate through the $(2n+1)$st order. This theorem is obvious, since an error of order $(n+1)$ in the wave function produces an error of order $(2n+2)$ in the expectation value of the energy. Thus, if we know the wave function accurate through the first order, we can calculate the energy accurate through the third order. For many chemical purposes this may be sufficient.

II. APPLICATIONS OF PERTURBATION THEORY

Perturbation theory can now be applied to a wide range of problems:

1) The calculation of molecular energies and the improvement of approximate eigenfunctions.

2) The determination of time-independent molecular properties.

3) The determination of properties of molecules in external fields.

4) The calculation of transition probabilities and off-diagonal matrix elements.

5) Effects of time-dependent perturbations.

6) The splitting of degenerate energy states due to both internal and external perturbations.

7) The derivation of "sum rules".

8) The derivation of variational principles associated with perturbation theory.

Since the inception of quantum mechanics, perturbation theory has been considered the appropriate tool for dealing with the effects of external fields, with the long range interaction between atoms and molecules, and with the small internal perturbations such as those that give rise to the fine and hyperfine structure of spectral lines. However, it is only recently that perturbation theory has been applied seriously to the fundamental problems of quantum chemistry which are concerned with the binding energy and with the structure and physical properties of the molecules. For such applications, the perturbations involve electron correlations and the perturbations are frequently large.

III. CONVERGENCE

A great deal of research has been done on the nature of the eigenvalues and the eigenfunctions of the perturbed quantum mechanical operators. Friedrichs, Rejto and others will discuss this interesting topic in considerable detail. For most of our molecular quantum mechanical applications, the basic theorem of Rellich [2] and the corollary of Kato [3] provide a justification for the use of perturbation theory [4].

RELLICH (1939) THEOREM:

If H_o is self-adjoint and if ϕ is any function in the domain of H_o, then the Rayleigh-Schrödinger perturbation series converge for $E(\lambda)$ and $\Psi(\lambda)$ for sufficiently small values of λ provided that two constants a and b exist such that

$$\langle V\phi | V\phi \rangle \leq a\langle H_o\phi | H_o\phi \rangle + b\langle \phi | \phi \rangle \ .$$

KATO (1951) COROLLARY:

The Rayleigh-Schrödinger perturbation theory applies (for sufficiently small λ) to any decomposition of the electrostatic Hamiltonian H for any atom, molecule, or finite crystal into two parts, H_o and λV,

$$H = H_o + \lambda V$$

provided that no new singularities stronger than Coulombic poles are introduced.

IV. THE BRILLOUIN-WIGNER AND RAYLEIGH-SCHRÖDINGER PERTURBATION EQUATIONS.

There are two principal types of perturbation theory which can be used: Rayleigh-Schrödinger and Brillouin-Wigner. The two procedures have much in common. In both, the Hamiltonian H for the perturbed system is split into H_o, the Hamiltonian for the unperturbed system, and λV, the perturbation potential:

$$H = H_o + \lambda V. \tag{1}$$

We fix our attention on the non-degenerate [5] energy state "o" which has the perturbed energy E_o and the wave function Ψ_o, so that

$$(H - E_o)\Psi_o = 0. \tag{2}$$

We assume that E_o and Ψ_o can be expressed as power series in the parameter λ, writing

$$E_o = \sum_{n=0}^{\infty} \lambda^n \epsilon_o^{(n)} \quad ; \quad \Psi_o = \sum_{n=0}^{\infty} \lambda^n \psi_o^{(n)} . \tag{3)-(4}$$

It is supposed that we know the wave function $\psi_o = \psi_o^{(o)}$ and its corresponding energy $\epsilon_o = \epsilon_o^{(o)}$. Thus,

$$(H_o - \epsilon_o)\psi_o = 0 . \tag{5}$$

The Rayleigh-Schrödinger and the Brillouin-Wigner procedures then differ in the manner in which they resolve the Schrödinger equation. [6]:

RAYLEIGH-SCHRÖDINGER

$$(H - E_o)\Psi_o = (H_o - \epsilon_o)\psi_o + \sum_{n=1}^{\infty} \lambda^n \left[\begin{array}{l} (H_o - \epsilon_o)\psi_o^{(n)} + V\psi_o^{(n-1)} \\ \qquad - \sum_{k=1}^{\infty} \epsilon_o^{(k)} \psi_o^{(n-k)} \end{array} \right] = 0 \tag{6}$$

BRILLOUIN-WIGNER

$$(7)\qquad (H-E_o)\Psi_o = (H_o-\epsilon_o)\psi_o + \sum_{n=1}^{\infty}\lambda^n[(H_o-E_o)\psi_o^{(n)} + V\psi_o^{(n-1)} - \epsilon_o^{(n)}\psi_o] = 0 \ .$$

If these equations are to remain valid as the perturbation parameter λ is varied over a continuous range of values, the coefficient of each power of λ must be individually zero. Thus, we obtain the Rayleigh-Schrödinger and the Brillouin-Wigner perturbation equations:

RAYLEIGH-SCHRÖDINGER

$$(8)\qquad (H_o-\epsilon_o)\psi_o^{(n)} = -V\psi_o^{(n-1)} + \sum_{k=1}^{n}\epsilon_o^{(k)}\psi_o^{(n-k)} \ .$$

BRILLOUIN-WIGNER

$$(9)\qquad (H_o-E_o)\psi_o^{(n)} = -V\psi_o^{(n-1)} + \epsilon_o^{(n)}\psi_o \ .$$

The more familiar forms of the Brillouin-Wigner equations may be obtained by expression $\psi_o^{(n)}$ and $\epsilon_o^{(n)}$ in terms of the complete set of eigenfunctions ψ_k and eigenvalues ϵ_k of the unperturbed Hamiltonian H_o. If we let $V_{jk} = \langle\psi_j|V|\psi_k\rangle$, then

$$\psi_o^{(n)} = \sum_{j_1}{}' \sum_{j_2}{}' \cdots \sum_{j_n}{}' \frac{\psi_{j_1} V_{j_1 j_2} V_{j_2 j_3} \cdots V_{j_n o}}{(E_o-\epsilon_{j_1})(E_o-\epsilon_{j_2})\cdots(E_o-\epsilon_{j_n})} \ ,$$

and

$$E_o^{(n)} = \sum_{j_1}{}' \sum_{j_2}{}' \cdots \sum_{j_n}{}' \frac{V_{oj_1} V_{j_1 j_2} \cdots V_{j_n o}}{(E_o-\epsilon_{j_1})(E_o-\epsilon_{j_2})\cdots(E_o-\epsilon_{j_n})} \ .$$

The prime on the summation sign indicates that none of the $j_1, j_2, \ldots, j_n$ should equal to "o".

If we do not know the complete set of unperturbed eigenfunctions and eigenvalues, the Brillouin-Wigner equations are much more difficult to solve than the Rayleigh-Schrödinger. First, we do not know a solution to the homogeneous equation $(H_o - E_o)\psi_o^{(n)} = 0$ of the Brillouin-Wigner Eq. (9). Second, for most practical problems, since E_o is not known in the beginning, trial values $\tilde{E}_o$ must be used in Eq. (9). Thus $\psi_o^{(n)}$ and $\epsilon_o^{(n)}$ are obtained as [7] functions of $\tilde{E}_o$. Finally, the correct value of E_o is determined by requiring the satisfaction of Eq. (3). Thus, the direct solution of the Brillouin-Wigner perturbation equations is unwieldy. As a result, very few of the new developments utilize the Brillouin-Wigner treatment.

V. DIRECT SOLUTION OF THE RAYLEIGH-SCHRÖDINGER PERTURBATION EQUATIONS

In contrast, the direct solution to the Rayleigh-Schrödinger Eq. (8) is more attractive. We know a solution ψ_o to the homogeneous equation $(H_o - \epsilon_o)\psi_o^{(n)} = 0$ corresponding to Eq. (8). As a result, we can use the method of Wronskians, Green's functions, or other well-known techniques to solve the inhomogeneous equation. For example, as Dalgarno and Lewis have suggested [8], if V is a real Hermitian operator, we can define a set of real scalar functions F_n such that $\psi_o^{(n)} = F_n\psi_o$. Then Eq. (8) becomes

$$[H_o, F_n]\psi_o = -VF_{n-1}\psi_o + \sum_{k=1}^{n} \epsilon_o^{(k)} F_{n-k}\psi_o \,. \tag{10}$$

Furthermore, if H_o has the form

$$H_o = -\frac{1}{2}\sum_j \nabla_j \cdot \nabla_j + U_o \,, \tag{11}$$

where U_o is a scalar function (which therefore commutes with F_n), then Eq. (10) can be expressed in the form

$$\sum_j \nabla_j \cdot (\psi_o^* \psi_o \nabla_j F_n) = \psi_o^* VF_{n-1}\psi_o + \psi_o VF_{n-1}\psi_o^* - 2\psi_o^*\psi_o \sum_{k=1}^{n} \epsilon_o^{(k)} F_{n-k} \,. \tag{12}$$

If Eq. (12) is one dimensional or separable, it may be integrated by quadratures.

In order to complete the specification of the functions F_n it is necessary to specify the normalization of the wave functions [9]. If we require that Ψ_o, as well as ψ_o, is normalized to unity, then it follows that

$$\sum_{k=0}^{n} \langle \psi_o^{(k)} | \psi_o^{(n-k)} \rangle = 0 \tag{13}$$

or

$$\sum_{k=0}^{n} \langle \psi_o | F_k F_{n-k} | \psi_o \rangle = 0 \ .$$

For a problem in one-dimensional cartesian coordinates,

$$F_n = \int_{-\infty}^{x} \frac{dx'}{\psi_o^* \psi_o} \int_{-\infty}^{x'} S_n(x'')\, dx'' \ . \tag{14}$$

Here $S_n(x)$ is the right hand side of Eq. (12). If ψ_o has nodes, then F_n can have poles at the nodal points [10]. If the nodal points are at $x = a_1, a_2, \ldots, a_\ell$ and we let

$$C_k = \left[\left(\frac{d\psi_o}{dx} \right)_{x=a_k} \right]^{-2} \int_{-\infty}^{a_k} S_n(x'')\, dx'' \ , \tag{15}$$

then Eq. (14) should be replaced by

$$F_n = \int_{-\infty}^{x} \frac{dx'}{\psi_o^* \psi_o} \left[- \sum_{k=1}^{\ell} \frac{C_k \psi_o^* \psi_o}{(x'-a_k)^2} + \int_{-\infty}^{x'} S_n(x'')\, dx'' \right] \ .$$

The same result may be obtained by replacing the integration over x' in Eq. (14) by a contour integral following any path between $-\infty$ and x which avoids the nodal points which occur on the real axis.

If Eq. (12) is not separable and ψ_o has complicated nodal surfaces, the integration of Eq. (12) may pose some interesting mathematical problems! Clearly, one would never use the Dalgarno-Lewis substitution $\psi_o^{(n)} = F_n \psi_o$ for such a case. Instead, Brueckner suggests that we express the n-th order perturbed function as

(16) $$\psi^{(n)} = [F_n + \nabla G_n \cdot \nabla]\psi_o$$

where the function G_n is chosen so as to properly shift the nodes of the wave function as a result of the perturbation. Eq. (16) takes advantage of the fact that the wave function and its gradient cannot simultaneously vanish. Brueckner and Gammel [11] have used essentially this form of $\psi^{(n)}$ in connection with the theory of finite nuclei.

VI. THE FIRST ORDER WAVE FUNCTION

If we know a zeroth order approximate wave function, we know the energy accurate through the first order. The first order perturbation equation may be written

(17) $$(H_o - \epsilon_o)\psi_o^{(1)} + (V - \epsilon_o^{(1)})\psi_o = 0 .$$

If we multiply Eq. (17) by ψ_o^* and integrate over all space, then since H_o is Hermitian, it follows that

$$\langle \psi_o | H_o - \epsilon_o | \psi_o^{(1)} \rangle = 0$$

and therefore,

(18) $$\epsilon_o^{(1)} = \langle \psi_o | V | \psi_o \rangle .$$

Indeed, the expectation value of the perturbed Hamiltonian corresponding to the zeroth order wave function is

(19) $$\mathcal{E}(0) = \langle \psi_o | H | \psi_o \rangle = \epsilon_o + \lambda\epsilon_o^{(1)} .$$

If now we solve Eq. (17) for $\psi_o^{(1)}$, then we can obtain the energy accurate through the third order. In order to see why this is true we need to consider the second and third order perturbation equations:

(20) $$(H_o - \epsilon_o)\psi_o^{(2)} + (V - \epsilon_o^{(1)})\psi_o^{(1)} = \epsilon_o^{(2)}\psi_o ,$$

(21) $$(H_o - \epsilon_o)\psi_o^{(3)} + (V - \epsilon_o^{(1)})\psi_o^{(2)} = \epsilon_o^{(2)}\psi_o^{(1)} + \epsilon_o^{(3)}\psi_o .$$

If we multiply Eq. (20) by ψ_o^* then, since

$$\langle \psi_o | H_o - \epsilon_o | \psi_o^{(2)} \rangle = 0 ,$$

it follows that

$$\epsilon_o^{(2)} = \langle \psi_o | V - \epsilon_o^{(1)} | \psi_o^{(1)} \rangle . \tag{22}$$

Similarly if we multiply Eq. (21) by ψ_o^* and integrate over all space we obtain

$$\epsilon_o^{(3)} = \langle \psi_o | V - \epsilon_o^{(1)} | \psi_o^{(2)} \rangle - \epsilon_o^{(2)} \langle \psi_o | \psi_o^{(1)} \rangle . \tag{23}$$

Now we can make use of the Hermitian properties of V and H_o, together with the normalization condition for $\psi_o^{(1)}$ and Eqs. (17) and (20), to perform the following chain of relations:

$$\begin{aligned} \epsilon_o^{(3)} &= \langle (V - \epsilon_o^{(1)}) \psi_o | \psi_o^{(2)} \rangle - \epsilon_o^{(2)} \langle \psi_o | \psi_o^{(1)} \rangle \\ &= - \langle (H_o - \epsilon_o) \psi_o^{(1)} | \psi_o^{(2)} \rangle - \epsilon_o^{(2)} \langle \psi_o | \psi_o^{(1)} \rangle \\ &= - \langle \psi_o^{(1)} | H_o - \epsilon_o | \psi_o^{(2)} \rangle - \epsilon_o^{(2)} \langle \psi_o | \psi_o^{(1)} \rangle \\ &= \langle \psi_o^{(1)} | V - \epsilon_o^{(1)} | \psi_o^{(1)} \rangle - \epsilon_o^{(2)} [\langle \psi_o^{(1)} | \psi_o \rangle + \langle \psi_o | \psi_o^{(1)} \rangle] \\ &= \langle \psi_o^{(1)} | V - \epsilon_o^{(1)} | \psi_o^{(1)} \rangle . \end{aligned} \tag{24}$$

It follows that the expectation value of the energy calculated with the wave function $\psi_o + \lambda \psi_o^{(1)}$ is given by

$$\mathcal{E}(1) = \epsilon_o + \lambda \epsilon_o^{(1)} + \frac{\lambda^2 \epsilon_o^{(2)} + \lambda^3 \epsilon_o^{(3)}}{1 + \lambda^2 \langle \psi_o^{(1)} | \psi_o^{(1)} \rangle} . \tag{25}$$

Eq. (25) is an example of the general Wigner theorem that a wave function accurate through the n-th order gives the energy accurate through the (2n + 1) st order.

From a formal standpoint, it is very easy to express $\psi_o^{(n)}$ and $\epsilon_o^{(n)}$ in terms of the resolvent

$$R_o = \frac{1 - |\psi_o><\psi_o|}{\epsilon_o - H_o} \quad .$$

Or, if the ψ_j and the ϵ_j form the complete set of eigenfunctions and eigenvalues of H_o, then the $\psi_o^{(n)}$ and $\epsilon_o^{(n)}$ can be expressed in terms of the equivalent spectral expansion. Thus:

$$\psi_o^{(1)} = R_o V \psi_o = \sum_j{}' \frac{V_{jo}\psi_j}{\epsilon_o - \epsilon_j} \tag{26}$$

$$\epsilon_o^{(2)} = <\psi_o | V R_o V | \psi_o> = \sum_j{}' \frac{V_{oj} V_{jo}}{\epsilon_o - \epsilon_j} \tag{27}$$

$$\epsilon_o^{(3)} = <\psi_o | V R_o (V - V_{oo}) R_o V | \psi_o> = \sum_j{}' \sum_k{}'' \frac{V_{oj} V_{jk} V_{ko}}{(\epsilon_o - \epsilon_j)(\epsilon_o - \epsilon_k)} . \tag{28}$$

Here

$$V_{jk} = <\psi_j | V | \psi_k> \ .$$

The summations correspond to a sum over all of the discrete states and an integral over the continuous energy states. The prime indicates that the state "o" is omitted. The double prime indicates that the states with $k = o$ and $k = j$ are omitted. From Eq. (27) it is easy to see that $\epsilon_o^{(2)}$ is necessarily zero or negative if the state "o" is the ground state.

We know the complete set of eigenfunctions and eigenvalues for the unperturbed hydrogen atom and for an unperturbed simple harmonic oscillator. Unfortunately, there are very few other practical cases where we could use such a spectral expansion. Thus, we seek explicit solutions to the perturbation equations.

In many perturbation problems we are given the Hamiltonian H together with an approximate wave function ψ_o. We are asked to determine, as best we can, the energy E and the wave function Ψ for the system. In order to put this problem in the framework of perturbation theory, it is necessary to determine a zeroth order Hamiltonian H_o. The operator H_o is neither unique nor obvious. Epstein will discuss the optimum determination of H_o. If ψ_o is a function of the spatial coordinates and does not involve spin, then we can write down the Sternheimer Hamiltonian,

(29) $$h_o = -\frac{1}{2}\sum_i \nabla_i \cdot \nabla_i + u_o$$

where

(30) $$u_o = \epsilon_o + \frac{1}{2}\sum_i (\nabla_i \cdot \nabla_i \psi_o)/\psi_o \ .$$

Thus, we have the obvious identity,

(31) $$h_o \psi_o = \epsilon_o \psi_o \ .$$

If we require that our zeroth order Hamiltonian have a local potential, then the zeroth order Hamiltonian H_o is uniquely equal to h_o.

If H is spin-free, then for a many electron problem, ψ_o can be written as a sum of spatial functions $\phi_j(\underline{r})$ multiplied by spin functions $\chi_j(\underline{s})$:

(32) $$\psi_o = \sum_j \phi_j(\underline{r})\,\chi_j(\underline{s}) \ .$$

Here the spin functions $\chi_j(\underline{s})$ are eigenfunctions of both the operator for the square of the spin and the z-component of the spin, corresponding to a particular spin state of the system. The $\chi_j(\underline{s})$ form the basis for an irreducible representation of the permutation group. The $\phi_j(\underline{r})$ should form the basis to the representation of the permutation group corresponding to the conjugate Young diagram [12]. The exact wave function is expressed in terms of the same spin functions, but different spatial functions:

(33) $$\Psi = \sum_j \Phi_j(\underline{r})\,\chi_j(\underline{s}) \ .$$

This suggests that we should be able to develop a spin-free perturbation expansion of a particular $\Phi_j(\underline{r})$ starting with the corresponding $\phi_j(r)$ as the zeroth order wave function. The corresponding Sternheimer potential is then

0a0

(34) $$u_o = \epsilon_o + \frac{1}{2}\sum_i (\nabla_i \cdot \nabla_i \phi_j)/\phi_j \ .$$

0b0

0c0

In most cases, the different $\phi_j(\underline{r})$ correspond to different zeroth order Sternheimer Hamiltonians H_o. As a result, it may be necessary to use projection operators to insure that the various orders of perturbation of a particular $\phi_j(\underline{r})$ have the symmetry with respect to permutations corresponding to the required Young diagram [13].

VII. NON-LOCAL POTENTIALS

It is possible to change the whole character of a perturbation problem by adding different types of non-local potentials to the Sternheimer Hamiltonian. For example, we could add "energy-shift" operators to form

$$H_o = h_o + \sum_k{}' a_k |\psi_k><\psi_k| \,. \tag{35}$$

Then if the ψ_k and ϵ_k are the eigenfunctions and eigenvalues of the Sternheimer Hamiltonian, $h_o\psi_k = \epsilon_k\psi_k$, the eigenfunctions of H_o remain unchanged but the energy values (except for the state "o") are shifted:

$$H_o\psi_o = \epsilon_o\psi_o \quad \text{and} \quad H_o\psi_j = (\epsilon_j + a_j)\psi_j \,. \tag{36}$$

If we use H_o as our zeroth order Hamiltonian, then from Eq. (26) is follows that

$$\psi_o^{(1)} = \sum_j{}' \frac{V_{jo}\psi_j}{(\epsilon_o - a_j) - \epsilon_j} \,. \tag{37}$$

Since ψ_j is orthogonal to ψ_o, the perturbation λV can be either equal to $H - H_o$ or $H - h_o$. If we choose all the constants a_j equal $\epsilon_o - E$, then our Rayleigh-Schrödinger perturbation equations become equivalent to the Brillouin-Wigner. However, in principle, it is possible for us to choose each of the a_j so that the exact wave function Ψ is equal to $\psi_o + \lambda\psi_o^{(1)}$.

Feenberg has suggested the use of "Change of Scale" operators. For the zeroth order Hamiltonian we take

$$H_o = h_o + \left(\frac{1-c}{c}\right)(h_o - \epsilon_o) \,; \tag{38}$$

where c is an arbitrary constant. Here again, the ψ_j remain

eigenfunctions of H_o. However,

$$H_o\psi_j = \frac{1}{c}[\epsilon_j - (1-c)\epsilon_o]\psi_j \quad . \tag{39}$$

Then it follows that

$$\psi_o^{(1)} = c \sum_j{}' \frac{V_{jo}V_j}{\epsilon_o - \epsilon_j} \quad . \tag{40}$$

Thus, the change of scale operators have the effect of multiplying the first order wave function by a constant. The constant c can be varied so as to optimize the expectation value of the energy calculated with the sum of the zeroth and first order functions. Dalgarno and Stewart [14] have used this variation to good advantage.

VIII. THE HYLLERAAS VARIATIONAL PRINCIPLE

If the first order perturbation equation is too difficult to solve exactly, we can make use of the Hylleraas variational principle to optimize an approximate first order wave function $\tilde{\psi}_o^{(1)}$. The Hylleraas principle states that for the ground state (or lowest energy state of a given symmetry)

$$\epsilon_o^{(2)} \leq \tilde{\epsilon}_o^{(2)} \quad , \tag{41}$$

where

$$\tilde{\epsilon}_o^{(2)} = \langle\tilde{\psi}_o^{(1)}|H_o - \epsilon_o|\tilde{\psi}_o^{(1)}\rangle + \langle\psi_o|V - \epsilon_o^{(1)}|\tilde{\psi}_o^{(1)}\rangle + \langle\tilde{\psi}_o^{(1)}|V - \epsilon_o^{(1)}|\psi_o\rangle \quad . \tag{42}$$

The proof of the Hylleraas principle is very simple. If $\tilde{E}$ is the expectation value of H calculated with the wave function $\psi_o + \lambda\tilde{\psi}_o^{(1)}$, then it is easy to show that

$$\tilde{E} = \epsilon_o + \lambda\epsilon_o^{(1)} + \lambda^2\tilde{\epsilon}_o^{(2)} + O(\lambda^3) \quad . \tag{43}$$

This is to be compared with the expansion of the exact energy

$$E = \epsilon_o + \lambda\epsilon_o^{(1)} + \lambda^2\epsilon_o^{(2)} + \lambda^3\epsilon_o^{(3)} + \cdots \quad . \tag{44}$$

By the Rayleigh-Ritz variational principle, if the state "o" is the lowest energy state of its symmetry, then $\tilde{E} \geq E$. The Hylleraas principle follows.

Sando and Hirschfelder [15] have generalized the Hylleraas principle so that if a wave function for a ground state (or lowest energy state of a given symmetry) is known through order (n-1), then an approximate n-th order function $\tilde{\psi}_o^{(n)}$ can be optimized by using the variational principle

$$\epsilon_o^{(2n)} \leq \tilde{\epsilon}_o^{(2n)} , \tag{45}$$

where (using the normalization condition of Eq. (12))

$$\begin{aligned} \tilde{\epsilon}_o^{(2n)} = {} & \langle \tilde{\psi}^{(n)} | H_o - \epsilon_o | \tilde{\psi}^{(n)} \rangle + \langle \tilde{\psi}_o^{(n)} | V - \epsilon_o^{(1)} | \psi_o^{(n-1)} \rangle \\ & + \langle \psi_o^{(n-1)} | V - \epsilon_o^{(1)} | \tilde{\psi}_o^{(n)} \rangle \\ & - \sum_{j=2}^{n-1} \epsilon^{(j)} [\langle \tilde{\psi}_o^{(n)} | \psi_o^{(n-j)} \rangle + \langle \psi_o^{(n-j)} | \tilde{\psi}_o^{(n)} \rangle] \\ & - \sum_{i=1}^{n-1} \sum_{j=n-i+1}^{n-1} \epsilon^{(j)} \langle \psi^{(i)} | \psi^{(2n-i-j)} \rangle . \end{aligned} \tag{46}$$

IX. THE ELECTROSTATIC ANALOGY

Prager and Hirschfelder [16] were impressed by the similarity of the first order perturbation equation

$$\nabla \cdot (\psi_o^2 \nabla F_1) = 2\psi_o (V - \epsilon_o^{(1)}) \psi_o \tag{47}$$

and the electrostatic equation

$$\nabla \cdot (\kappa \nabla \phi) = - 4\pi\rho . \tag{48}$$

To make the analogy: F_1 corresponds to the electrostatic potential ϕ; the ψ_o^2 corresponds to the dielectric constant κ; also $-(4\pi)^{-1}\epsilon_o^{(2)}$ corresponds to the electrostatic self-energy of the charge distribution $U = \frac{1}{2}\int \rho\phi d\tau$; and $-\nabla F_1$ corresponds to the electric field strength $\mathcal{E} = -\nabla\phi$. Thus, we can use in perturbation theory the Thomson

and Dirichlet variational principles which were originally derived for electrostatics:

THOMSON'S PRINCIPLE states that if $\tilde{\mathcal{E}}$ is an approximate electric field vector subject to the condition $\nabla\cdot(\kappa\tilde{\mathcal{E}}) = 4\pi\rho$, then

$$U \leq \frac{1}{8\pi}\int \kappa\tilde{\mathcal{E}}\cdot\tilde{\mathcal{E}}\,d\tau \quad . \tag{49}$$

The analogous theorem states that if $\tilde{G}$ is a trial vector subject to the condition $\nabla\cdot(\psi_o^2\tilde{G}) = -2\psi_o(V-\epsilon_o^{(1)})\psi_o$, then

$$\epsilon_o^{(2)} \geq -\frac{1}{2}\int \psi_o^2\tilde{G}\cdot\tilde{G}\,d\tau \quad . \tag{50}$$

DIRICHLET'S PRINCIPLE states that if $\kappa\tilde{\phi}\nabla\tilde{\phi}$ or $\psi_o^2\tilde{F}\nabla\tilde{F}$ approaches zero faster than the reciprocal of the square of the distance in the limit as the distance approaches infinity, then

$$U \geq 2\pi\left[\int \rho\tilde{\phi}\,d\tau\right]^2 / \int \kappa\nabla\tilde{\phi}\cdot\nabla\tilde{\phi}\,d\tau \quad . \tag{51}$$

Or

$$\epsilon_o^{(2)} \leq -2\left[\int \psi_o(V-\epsilon_o^{(1)})\tilde{F}\psi_o\,d\tau\right]^2 / \int \psi_o^2\nabla\tilde{F}\cdot\nabla\tilde{F}\,d\tau \quad . \tag{52}$$

Actually, the Dirichlet and Hylleraas principles are equivalent. If we set $\tilde{\psi}_o^{(1)} = a\tilde{F}\psi_o$ in the Hylleraas principle, Eqs. (41) and (42), and optimize the constant a, then we obtain the Dirichlet principle, Eq. (52).

Thus, the Thomson principle gives a lower bound and the Dirichlet (or Hylleraas) principle gives an upper bound to the second order energy. These principles are particularly useful in the calculation of the polarizability. If an atom or molecule is placed in a constant electric field, the polarizability $\alpha = -2\epsilon_o^{(2)}$ when the electric field strength is taken as the perturbation parameter. Let us consider a simple example:

EXAMPLE: Polarizability of 1s Atomic Hydrogen

$$\psi_o = (\pi)^{-1/2}\exp(-r); \quad \lambda = \mathcal{E}_x$$

$$V = -x; \quad \epsilon_o^{(1)} = 0 \quad .$$

Thomson Principle:

$$\epsilon_o^{(2)} \geq -\frac{1}{2\pi}\int e^{-2r}\tilde{G}\cdot\tilde{G}\,d\tau$$

subject to the condition

$$\nabla\cdot[e^{-2r}\tilde{G}] = 2xe^{-2r} .$$

Let us assume the trial field vector

$\tilde{G} = (r + \frac{1}{2})$ times a unit vector in the x-direction.

The Thomson principle then gives $\epsilon_o^{(2)} \geq -2.375$ or $\alpha \leq 4.75$.

Dirichlet or Hylleraas Principle: Let us assume that $\tilde{F} = ax$. Then,

$$\epsilon_o^{(2)} \leq \frac{2}{\pi}\int x^2 e^{-2r}\,d\tau = -2 .$$

Thus, $\alpha \geq 4.00$.

The Thomson and Dirichlet principles have shown for this example that $4.00 \leq \alpha \leq 4.75$. The correct value is $\alpha = 4.50$. If we had used more elaborate trial functions, it would have been possible to obtain very accurate upper and lower bounds for the polarizability. Of course, this example is especially favorable since we know the exact solution for the system without the external field. More usually, we only know approximate solutions to the zero field problem and more complicated double perturbation procedures are required in order to determine the polarizability (see Ref. 26).

X. THE EVALUATION OF INFINITE SUMS

Dalgarno has been very successful in using the techniques of perturbation theory in the evaluation of infinite sums. For example, α_{xx}, the xx-component of the polarizability, can be expressed in terms of μ_x, the x-component of the dipole moment, by the summation

$$\alpha_{xx} = 2\sum_j{}' \frac{(\mu_x)_{oj}(\mu_x)_{jo}}{\epsilon_j - \epsilon_o} . \qquad (53)$$

If we could find a function F for which the matrix component F_{jo} satisfies the relation

(54) $$F_{jo} = (\mu_x)_{jo}/(\epsilon_j - \epsilon_o) ,$$

then we could use the rules of matrix multiplication to sum the series,

(55) $$\alpha_{xx} = 2\sum_j{}' (\mu_x)_{oj} F_{jo} = 2(\mu_x F)_{oo} - 2(\mu_x)_{oo} F_{oo} .$$

The term with the minus sign is due to the fact that $j = o$ is not included in the sum. In order to determine the function F it is necessary to solve the differential equation

(56) $$(H_o - \epsilon_o)F\psi_o = [\mu_x - (\mu_x)_{oo}]\psi_o .$$

If we multiply Eq. (56) by ψ_j^* and integrate over all space, it is obvious that F_{jo} satisfies Eq. (54). Thus, the difficulty of evaluating the infinite summation has been replaced by the difficulty of solving the differential equation.

XI. CALCULATION OF ENERGY AND IMPROVEMENT OF APPROXIMATE WAVE FUNCTIONS

One of the principal applications of perturbation theory is the calculation of the energy and the improvement of the wave function. We are given the Hamiltonian H for the system. We are also given a function ψ_o which we are told to use as a zeroth order approximation to the eigenfunction of some state of the system. There are four classes of examples which have been studied extensively:

1) Expansion in powers of Z^{-1} for atoms and diatomic molecules. Here Z is the atomic number.

2) The approximation of, and improvement of, Hartree-Fock wave functions and energy.

3) Expansion in powers of R^{-1} for long range intermolecular potentials. Here R is the separation between the molecules.

4) The "united atom" expansion of molecular energy in powers of R.

In the early days of quantum mechanics, Hylleraas showed that if the Schrödinger equation for an atom is divided by Z^2 and the unit of length is changed, so that $r' = Zr$ and the scale of energy is changed so that $E' = Z^{-2}E$, then we get a new Schrödinger equation $H'\Psi(r') = E'\Psi(r')$ where

$$(57) \qquad H' = \sum_i [-\tfrac{1}{2}\Delta_i' - (r_i')^{-1}] + Z^{-1} \sum_{i<j} (r_{ij}')^{-1} .$$

Thus, Z^{-1} forms a natural perturbation parameter. The atomic hydrogen orbitals furnish the zeroth order wave function. The Rayleigh-Schrödinger perturbation sequence seems empirically to converge for $Z > 0.78$ for two-electron atoms or ions. Charles Scherr will tell us about his elegant work with Robert Knight on the Z^{-1} expansions of atomic energy states. Matcha and Byers Brown [17] have used a Z^{-1} expansion for considering the hydrogen molecule. For homonuclear diatomic molecules, the Z^{-1} expansion is carried out with $\rho = RZ$ held constant. Cohen [18] has used the Z^{-1} expansion to approximate Hartree-Fock orbitals. Even low order orbitals give good approximations to the values for the Hartree-Fock energy of atoms. Perturbation theory can also be used to calculate the correlation energy which is missing from the Hartree-Fock treatment.

Also, since the early days of quantum mechanics, long range intermolecular potentials have been expanded in powers of R^{-1} where R is the separation between molecules A and B. The Hamiltonian for the two molecule system is expressed as the sum of the Hamiltonian of A plus the Hamiltonian of B plus an interaction potential V_{ab}. As long as the charge distributions of A and B do not overlap, it is not necessary to consider the exchange of electrons between A and B. If almost all of the electronic charge density of molecule A lies within a radius r_a and, similarly, almost all of the charge density of molecule B lies within a radius r_b, then if $R > r_a + r_b$, the interaction potential V_{ab} can be expanded in powers of R^{-1}. For the interaction of two neutral molecules, the leading term in the expansion of V_{ab} is $O(R^{-3})$ and the usual London dispersion energy, $-C_{ab}/R^6$, is given by the second order perturbation energy. Dalgarno and others are calculating very precise values of C_{ab} which can be used for a variety of experimental applications. However, the overlapping of the charge distributions in the two molecules leads to terms which vary as $\exp(-cR)$. Since $\exp(-cR)$ is not an analytic function of R^{-1}, it follows that R^{-1} is <u>not</u> a natural perturbation parameter and expansions in powers of R^{-1} have only limited validity.

Buckingham, Bingel, and Byers Brown and Steiner have considered united atom expansions of the molecular energy. Here the Hamiltonian is expanded in powers of R starting with the united atom corresponding to a confluence of all of the nuclei. Such an expansion necessarily neglects those regions of electron configuration space where the electrons lie between the nuclei. As a result, Byers Brown and Steiner [19] have found that the energy expansion contains a term of the order of $R^5 \log R$. Hence R is not a natural

perturbation parameter. However, there still remains the possibility that the united atom expansions will be very useful.

We hope that for many chemical purposes sufficient accuracy will be obtained by calculating the approximate wave function perturbed through the first order and the energy accurate through the third order. The recent calculations for the ground state of H_2^+ by Lyon, Matcha, Sanders, Meath, and Hirschfelder [20] have shown that starting with the simplest functional forms for the zeroth order wave function, the wave function through the first order is accurate to one part in 10,000 for most nuclear and electronic configurations. The energy through the third order is accurate to within $0.0001\ e^2/a_o$ which corresponds to 0.06 kcal/mole. Similarly, Matcha and Byers Brown [17] have started with the diatomic hydrogen ion wave function as a zeroth order orbital for the ground state of H_2 and have obtained the energy through third order accurate to within $0.0003\ e^2/a_o$ or 0.2 kcal/mole.

Most of the approximate wave functions which are used in perturbation theory do not satisfy the cusp conditions [21] so that in the limit as two particles come together, $H\psi_o$ becomes unbounded. Conroy [22] has shown how to construct approximate molecular wave functions which satisfy all of the boundary conditions and behave properly at the singular points in the potential. From the standpoint of strictly variational calculations, the behavior at the singular points seems to be unimportant. However, in perturbation calculations, correct behavior of ψ_o near the singular points may make a large difference in the accuracy of the results.

XII. EXPECTATION VALUES OF PROPERTIES OTHER THAN ENERGY

The calculation of the expectation values of properties other than energy can be treated within the framework of perturbation theory. We distinguish two types of properties: first order, such as the dipole moment; and second order, such as polarizability. For both types of properties it is expedient to consider a real or fictitious Hamiltonian

$$\mathcal{H} = H + \mu W \ . \tag{58}$$

Here μW is the "external perturbation" where W is related to the property under consideration. Then the energy for the externally perturbed system is

$$\mathcal{E} = E + \mu E^{(1)} + \mu^2 E^2 + \ldots \ . \tag{59}$$

The system without the external perturbation satisfies the Schrödinger equation

(60) $$H\Psi = E\Psi \ .$$

The expectation values of first order properties can be expressed in the form

(61) $$\langle W \rangle = \langle \Psi | W | \Psi \rangle = E^{(1)} \ .$$

The expectation values of second order properties are given by

(62) $$\langle Q \rangle = \langle \Psi | W - \langle W \rangle | \Psi^{(1)} \rangle = E^{(2)} \ .$$

From a formal standpoint, $\langle Q \rangle$ is the first order expectation value of the symbolic operator $Q = -(W - \langle W \rangle)(H - E)^{-1}(W - \langle W \rangle)$.

Unfortunately we seldom know the exact wave function Ψ or its energy E. Instead, we know an approximate wave function ψ_o which satisfies the Schrödinger equation $H_o \psi_o = \epsilon_o \psi_o$. We can define the "internal perturbation" which results from the "badness" of the approximate wave function as $\lambda V = H - H_o$. Then, since $\Psi = \psi_o + \lambda \psi_o^{(1)} + \lambda^2 \psi_o^{(2)} + \ldots$, we can express $\langle W \rangle$ as a power series in λ,

(63) $$\langle W \rangle = \langle W \rangle_o + \lambda \langle W \rangle_1 + \ldots \ ,$$

where

(64) $$\langle W \rangle_o = \langle \psi_o | W | \psi_o \rangle \ ,$$

and $\langle W \rangle_1$, the first correction for the "badness" of the approximate wave function, is given by

(65) $$\langle W \rangle_1 = \langle \psi_o^{(1)} | W | \psi_o \rangle + \langle \psi_o | W | \psi_o^{(1)} \rangle \ .$$

Or, since $\langle \psi_o^{(1)} | \psi_o \rangle + \langle \psi_o | \psi_o^{(1)} \rangle = 0$,

(66) $$\langle W \rangle_1 = \langle \psi_o^{(1)} | W - W_{oo} | \psi_o \rangle + \langle \psi_o | W - W_{oo} | \psi_o^{(1)} \rangle \ .$$

Here $\psi_o^{(1)}$ is the solution to the first order equation

(67) $$(H_o - \epsilon_o)\psi_o^{(1)} = -(V - V_{oo})\psi_o \ .$$

Generally, the internal perturbation V involves $1/r_{ij}$ electron correlation terms and Eq. (67) may be very difficult to solve. On

this account, we are fortunate to have the Dalgarno Interchange Theorem which states that

$$(68) \qquad \langle W\rangle_1 = \langle \psi_o^{(o,1)} | V - V_{oo} | \psi_o \rangle + \langle \psi_o | V - V_{oo} | \psi_o^{(o,1)} \rangle ,$$

where $\psi_o^{(o,1)}$ is the first order perturbed wave function corresponding to the external perturbation and satisfies the equation

$$(69) \qquad (H_o - \epsilon_o)\psi_o^{(o,1)} = -(W - W_{oo})\psi_o .$$

The operators W are usually one-electron operators so that Eq. (69) is frequently separable and its solutions may even be expressible in closed form. In any case, Eq. (69) is usually much easier to solve than Eq. (67). Thus, $\langle W\rangle_1$ is usually easy to calculate. Unfortunately there is no interchange theorem which applies to $\langle W\rangle_2$ and $\langle W\rangle_2$ is generally difficult to calculate. The interchange theorem is easy to prove. Multiply Eq. (69) by $\psi_o^{(1)*}$ and integrate and compare with the result of multiplying Eq. (69) by $\psi_o^{(o,1)*}$ and integrating. Thus,

$$(70) \qquad \begin{aligned} \langle \psi_o^{(1)} | H_o - \epsilon_o | \psi_o^{(o,1)} \rangle &= -\langle \psi_o^{(1)} | W - W_{oo} | \psi_o \rangle \\ &= \langle \psi_o^{(o,1)} | H_o - \epsilon_o | \psi_o^{(1)} \rangle^* \\ &= -\langle \psi_o^{(o,1)} | V - V_{oo} | \psi_o \rangle^* \\ &= -\langle \psi_o | V - V_{oo} | \psi_o^{(o,1)} \rangle . \end{aligned}$$

Löwdin points out that the interchange theorem is obvious if one expresses $\psi_o^{(1)} = R_o V \psi_o$ and $\psi_o^{(o,1)} = R_o W \psi_o$, where R_o is the symbolic operator

$$(71) \qquad R_o = \frac{1 - |\psi_o\rangle\langle\psi_o|}{\epsilon_o - H_o} ,$$

so that

$$(72) \qquad \langle W\rangle_1 = \langle \psi_o | V R_o W + W R_o V | \psi_o \rangle .$$

The interchange theorem corresponds to the symmetry between V and W.

Dalgarno made the further hypothesis or conjecture that if ψ_o contains an embedded parameter which can be varied so as to make $\langle W\rangle_1 = 0$, then the value of $\langle W\rangle_o$ corresponding to this value of the parameter is a better approximation to $\langle W\rangle$ than is $\langle W\rangle_o + \lambda\langle W\rangle_1$ calculated using a different value of the parameter. The Dalgarno hypothesis gives good numerical results in a large number of cases. Robinson [23] gives the following explanation. If $\langle W\rangle_1 = 0$, then ψ_o satisfies the hypervirial theorem [24]

$$\langle \psi_o | [H, L] | \psi_o \rangle = 0 \tag{73}$$

where L is an anti-Hermitian operator satisfying the condition $\psi_o^{(o,1)} = L\psi_o$. The satisfaction of the hypervirial theorem has the interesting consequence that the wave function ψ_o is energetically stable with respect to variations of the type [24] $\psi_o \rightarrow \psi_o + \mu\psi_o^{(o,1)}$, that is, the energetically optimum value of μ is zero. Thus, we can associate $\langle W\rangle_1 = 0$ with a variational principle for ψ_o.

Sanders and Hirschfelder [25] calculated the expectation values of the nth power of the radius of an electron in the ground state of helium taking $\psi_o = s^3(\pi)^{-1} \exp(-s(r_1 + r_2))$. Here n went from -2 to +6 and the constant s was varied. When s was adjusted to make $\langle r_1^n\rangle_1 = 0$, then the resulting values of $\langle r_1^n\rangle_o$ are comparable in accuracy with the Hartree-Fock values and approximately 91% of the correct value. For other choices of s, such as s energy-optimized, the values of $\langle r_1^n\rangle_o + \lambda\langle r_1^n\rangle_1$ are not as good.

In a similar manner, the expectation values of second order properties may be expanded in powers of λ. Thus,

$$\langle Q\rangle = \langle Q\rangle_o + \lambda\langle Q\rangle_1 + \ldots, \tag{74}$$

where

$$\langle Q\rangle_o = \langle \psi^{(o,1)} | W | \psi_o \rangle \tag{75}$$

and, using another Dalgarno interchange theorem,

$$\langle Q\rangle_1 = \langle \psi_o^{(o,2)} | V | \psi_o \rangle + \langle \psi_o^{(o,1)} | V | \psi_o^{(o,1)} \rangle + \langle \psi_o | V | \psi_o^{(o,2)} \rangle . \tag{76}$$

0a0

[illegible] wave function corresponding

0b0

0c0

Here $\psi^{(o,2)}$ is the second order perturbed wave function corresponding to the external perturbation. Since Eq. (76) involves only the externally perturbed wave functions, $<Q>_1$ is comparatively easy to evaluate. Again Dalgarno conjectures that if ψ_o contains an embedded parameter which can be varied so as to make $<Q>_1 = 0$, then the corresponding value of $<Q>_o$ is the best estimate of $<Q>$. And again numerical calculations support his hypothesis.

There are two ways that Hartree-Fock wave functions may be used to calculate expectation values: the coupled and the uncoupled Hartree-Fock approximations [26]. The coupled Hartree-Fock provides more accurate results but it is much more difficult to carry out, since it requires obtaining a new set of Hartree-Fock orbitals corresponding to the Hamiltonian $\mathcal{H}$. In the coupled Hartree-Fock, $<W>_1 = 0$ and $<Q>_1 = 0$. In the uncoupled Hartree-Fock, $<W>_1 = 0$ but $<Q>_1$ is not equal to zero. Recently Tuan, Epstein, and Hirschfelder [27] showed that (thanks to the Brillouin theorem) the uncoupled Hartree-Fock $<Q>_1$ can be expressed rather simply in terms of the first order externally perturbed orbitals which are used in the calculation of $<Q>_o$. Thus, we can correct the uncoupled Hartree-Fock expressions for $<Q>_o$ for the "badness" of the wave function.

XIII. MATHEMATICAL PROBLEMS

There are basically two types of mathematical problems associated with the use of perturbation theory:

1. First, there is the question as to how the perturbation changes the eigenvalue spectrum of the Hamiltonian. For example, if a molecule is placed in an electric field of constant field strength (no matter how small), all of the discrete energy levels are shifted and at the same time broadened into dense packets of continuous energy levels which physicists call "metastable" states. This is illustrated in Fig. 1. The Auger effect and the Lamb shift for excited state atoms furnish other examples where the discrete energy eigenvalues of H_o "disappear into the continuum" with the application of the perturbation. Fortunately, Kato [3] and Titchmarsh [28] have studied such problems and have provided us with criteria for H_o and V such that the average energy of the metastable state is correctly calculated by perturbation theory.

2. Second, there is the question of the convergence of the perturbation series. Does the perturbation series converge in the limit as λ approaches zero? If so, what is the range of values of λ for which the perturbation series converges? Kato [3] has shown, by using a variational treatment, that even if the perturbation series

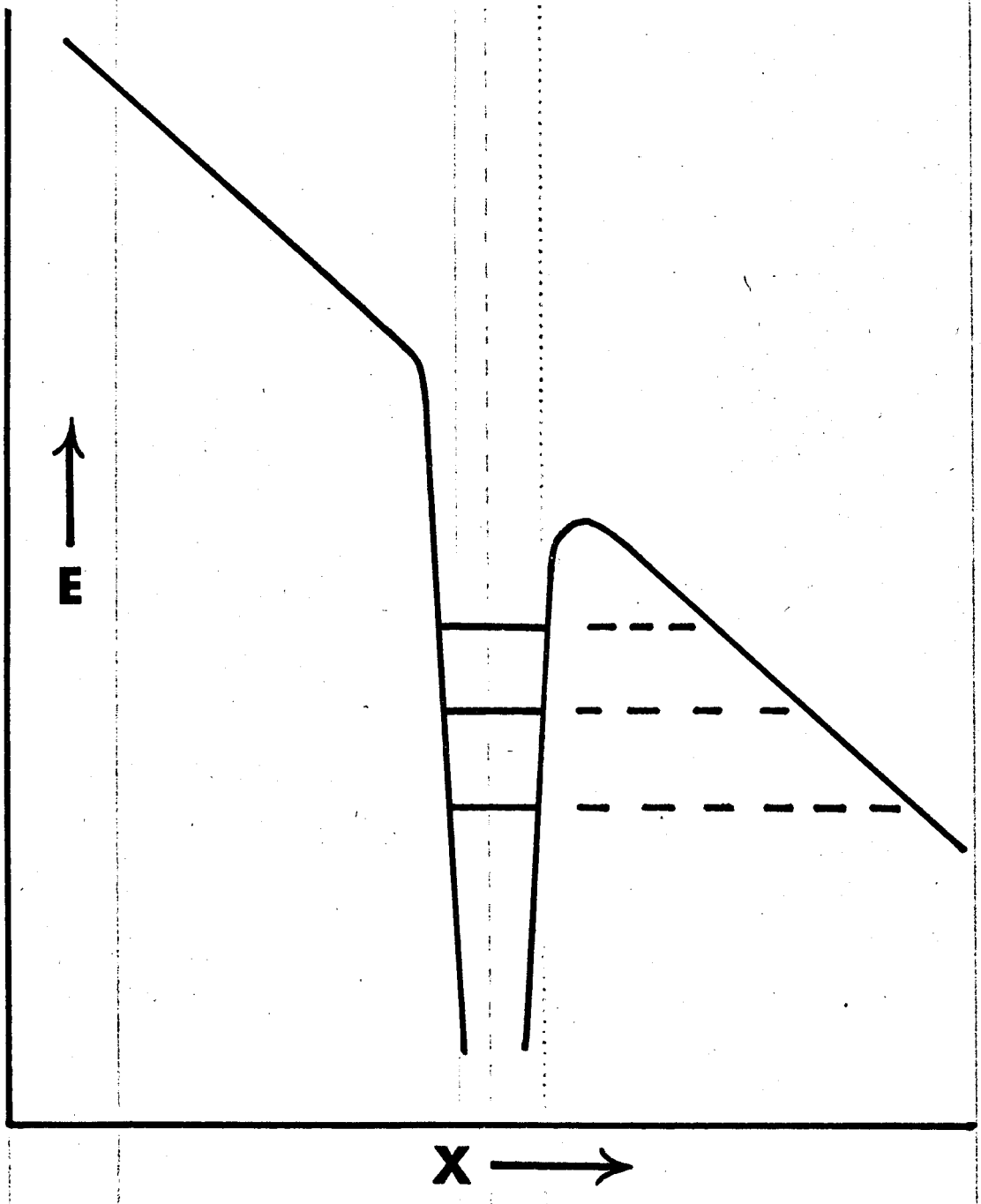

Fig. 1. The Electronic Potential Energy and the Energy Levels of a Hydrogen Atom Placed in an Electric Field. Since the electron can "tunnel" through the potential barrier, the perturbation of the electric field shifts the discrete hydrogen energy levels and broadens them into packets of continuous energy states.

only converges in an asymptotic sense, there is the possibility (depending upon the nature of H_o and λV) that the first few terms of the perturbation series (if they exist) may provide a useful approximation to the properties of the perturbed system. Kato [3] explains that "roughly speaking, perturbation theory gives correct results in the sense of asymptotic expansions as long as the necessary quantities are calculated by operations in Hilbert space".

As Friedrichs, McLeod, Conley, Brownell, Rejto, Kuroda, de Branges, and Phillips will tell you in this symposium, the nature of the spectrum and the nature of the convergence depend sensitively on the nature of the operators H_o and V. In molecular quantum mechanics, we are usually given an operator H and we are given great latitude in how H is resolved into H_o and λV. Thus, the answer to the question, "What is H_o?" will depend upon very general mathematical studies of the foundations of perturbation theory. There are three types of perturbed Hamiltonians which occur in quantum chemistry:

1. In calculating the energy of a molecule, we usually use an electrostatic Hamiltonian in which the only singularities are poles corresponding to the confluence of any pair of particles. The work of Kato [3] has given us justification for treating molecular energy problems within the framework of perturbation theory.

2. In calculating the hyperfine structure of a molecule, we sometimes use the Breit-Pauli Hamiltonian which contains singularities of the third order. What is the spectrum of such a Hamiltonian? Can we legitimately use the Breit-Pauli Hamiltonian to calculate the energy through the first order? Löwdin [29] has shown that these third order singularities do not occur in the original Dirac equations and therefore should not occur in the physical problem.

3. In calculating the expectation values of properties other than energy, we use as the perturbed Hamiltonian $\mathcal{H} = H + \mu W$. Here the external perturbation μW may correspond to a physically realizable external field or it may be a convenient mathematical fiction. In either case, W is the operator associated with the particular expectation value which we desire. The operator W may correspond to <u>any</u> physical property of the system. Adding μW to H can therefore lead to a very strange sort of Hamiltonian $\mathcal{H}$ with very strange spectrum, etc. What restrictions must be placed on the operators W in order that the eigenvalues of $\mathcal{H}$ are analytic functions of μ?

Other mathematical problems are less profound, but still puzzling. For example, we would like to know how to treat electron exchange as a perturbation. It is an unusual type of perturbation which is associated with an unusual type of degeneracy [30].

Consider the hydrogen molecule as an illustration [31]. We distinguish two zeroth order wave functions: $\psi_1 = a_1 b_2$, corresponding to electron 1 on atom A and electron 2 on atom B, and $\psi_2 = a_2 b_1$, which is the same function with the electrons interchanged. The expectation value of H calculated with ψ_2 gives the same energy as if it were calculated with ψ_1. This is the sense in which the two electron configurations are degenerate. However, ψ_1 and ψ_2 correspond to different H_o's and therefore they are not degenerate in the usual sense. Of course one can use a linear combination of ψ_1 and ψ_2 (with the proper symmetry) as a zeroth order function and proceed to form a standard perturbation calculation. However, intuitively one feels that it should be possible to use ψ_1 as the zeroth order function and express the effects of the required symmetrization or antisymmetrization as a perturbation operator. Primas [32] feels that the difficulty in doing this represents a defect in the formulation of quantum mechanics.

Another type of problem which requires further investigation is the perturbation treatment of almost-degenerate states. If these states are treated as non-degenerate, there are no formal difficulties but the perturbation series only converges for very small values of λ. Instead, we can consider explicitly the set of states whose energies lie close together and carry out a combined perturbation-variation (DEFOPVIM) treatment [1] in which neither the energy nor the wave function are analytic functions of λ, but the matrix components in the secular equation are analytic functions of λ. Such a scheme represents a generalization of our usual definition of perturbation theory. It has the practical advantage that it permits us to extend the domain of acceptable values of λ.

Indeed, we are interested in the general question of how we can improve the convergence of the perturbation series. For some types of problems, the convergence of the Brillouin-Wigner or Feenberg series is faster than for the Rayleigh-Schrödinger. However, for many-body problems,the Rayleigh-Schrödinger is required. Are there convenient criteria that we might use to tell us which series should be used ?

In many physical problems it is not clear what we should use as a perturbation parameter [33]. Actually, in the atomic energy problem, Z^{-1} did not appear to be the natural parameter until after we had changed the scale of our coordinates and energy. Transformation of variables may completely change the complexion of a perturbation problem and suggest a different choice of expansion parameter. Thus there is a search underway for natural perturbation parameters to use for long range intermolecular forces (where R^{-1} is not satisfactory); for the united atom expansion (where R is not satisfactory); etc.

One of the problems which has intrigued me, is the search for a fast converging perturbation scheme. The Rayleigh-Schrödinger formalism is needlessly slow in converging because it does not take advantage of the fact that after you have obtained the n-th order wave function, you know the energy accurate through the (2n + 1)st order. Thus, two years ago, I suggested [34] an iteration process (FOPIM) in which the sum of the zeroth and first order wave functions is used as the zeroth order wave function in a new calculation to obtain an improved first order function, etc. After the n-th iteration, the expectation value of the energy is given by a complicated non-analytic function of λ, which may be expanded in an asymptotic series in powers of λ accurate up to terms of the order of λ raised to the 2^{n+1} power. Unfortunately, the iterated perturbation potential is not analytic in λ and convergence difficulties make this method unattractive [35]. There are, however, many other types of iterative procedures which might be more practical [36, 37].

Unfortunately I did not have time to discuss the important and difficult problems of time-evolution, scattering, and many-body systems which are of great interest to theoretical chemists.

I have tried to explain how we are trying to determine the steady-state properties of molecules and what mathematical difficulties we are encountering. The better we can understand the basic mathematical nature of our equations and the mathematical structure of their solutions, the easier it will become to devise practical methods for determining the molecular properties.

FOOTNOTES AND REFERENCES

1. My talk is an outgrowth of a review article, Recent Developments in Perturbation Theory, by J. O. Hirschfelder, W. Byers Brown and S. T. Epstein which was published in Advances in Quantum Chemistry, Vol. I (Academic Press, New York, 1964; P. O. Löwdin, Editor). Indeed many of the ideas which I express are the result of very close collaboration and frequent consultation with Epstein and Byers Brown.

2. F. Rellich, Math. Ann. (Leipzig) 116, 555 (1939) and 117, 356 (1940).

3. T. Kato, J. Fac. Sci. Univ. Tokyo, Sec. I, 6, 145 (1951).

4. T. W. Palmer, Univ. Wisconsin Theoretical Chemistry Institute Report WIS-TCI-62, August (1964); see also Ref. (1), p. 349.

5. The treatment of degenerate states starts with the choice of the zeroth order wave functions which are linear combinations of the original unperturbed wave functions which diagonalize the perturbation matrix. The second order energy then has the same structure as for a non-degenerate problem. However, the first order wave function, the third order energy, etc. have additional terms as a result of the degeneracy. See A. Dalgarno, Quantum Theory (D. R. Bates, ed.) Vol. 1, Chapter 5, p. 194, Academic Press (New York), 1964. See also Ref. 1, p. 312. The proper treatment of degenerate problems does not seem to be unique.

6. W. Byers Brown and W. J. Meath, Proc. Nat. Acad. Sci. (U.S.A.), 52, 65 (1964).

7. W. J. Meath and J. O. Hirschfelder, J. Chem. Phys. 41, 1628 (1964).

8. A. Dalgarno and J. T. Lewis, Proc. Roy. Soc. (London) A233, 70 (1955).

9. For some problems it is convenient to use the "intermediate normalization" such that $\langle\psi_o|\psi_o^{(n)}\rangle = \langle\psi_o|F_n|\psi_o\rangle = 0$. For the intermediate normalization, Ψ is not normalized to unity.

10. W. Byers Brown and J. O. Hirschfelder, Proc. Nat. Acad. Sci. (U.S.A.) 50, 399 (1963).

11. See for example, K. A. Brueckner and J. L. Gammel, Phys. Rev. 109, 1021 (1958).

12. See for example, M. Hamermesh, "Group Theory" (Addison-Wesley, 1962), p. 422.

13. This procedure is equivalent to that of S. T. Epstein and J. H. Karl, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-90 (April, 1965).

14. A. Dalgarno and A. L. Stewart, Proc. Phys. Soc. (London) 77, 467 (1961).

15. K. M. Sando and J. O. Hirschfelder, Proc. Nat. Acad. Sci. (U.S.A.) 52, 434 (1964). See also R. E. Knight and C. W. Scherr, Revs. Mod. Phys. 35, 436 (1963), who have derived similar principles for the even-order energies through the twelfth-order.

16. S. Prager and J. O. Hirschfelder, J. Chem. Phys. 39, 3289 (1963).

17. R. L. Matcha and W. Byers Brown, J. Chem. Phys. 00, 0000 (1966). Also, R. L. Matcha, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-129 (October, 1965).

18. M. Cohen, Proc. Phys. Soc. (London) 82, 778 (1963).

19. W. Byers Brown and E. Steiner, University of Wisconsin Theoretical Institute Report WIS-TCI-000 (1965).

20. W. D. Lyon, R. L. Matcha, W. A. Sanders, W. J. Meath, and J. O. Hirschfelder, J. Chem. Phys. 43, 1095 (1965).

21. R. T. Pack and J. O. Hirschfelder, University of Wisconsin Theoretical Chemistry Institute Report WIS-TCI-123, November (1965).

22. H. Conroy, J. Chem. Phys. 41, 1327-1351 (1964).

23. P. D. Robinson, Proc. Roy. Soc. A283, 229 (1965).

24. S. T. Epstein and J. O. Hirschfelder, Phys. Rev. 123, 1495 (1961).

25. W. A. Sanders and J. O. Hirschfelder, J. Chem. Phys. 42, 2904 (1965) and 43, 0000 (Mulliken Issue, 1965). The desirability of requiring that $\langle W \rangle_1 = 0$ or that $\langle Q \rangle_1 = 0$ is probably restricted to operators W which are positive definite. The question whether $\langle W \rangle_0 + \lambda \langle W \rangle_1$ is stationary for embedded parameters chosen to make $\langle W \rangle_1 = 0$ depends upon the functional form of the zeroth order wave function.

26. A Dalgarno, Advances in Physics, 11, 281 (1962).

27. D. F. Tuan, S. T. Epstein, and J. O. Hirschfelder, J. Chem. Phys. 44, 0000, January 15 (1966).

28. E. C. Titchmarsh, "Eigenfunction Expansions Associated with Second Order Differential Equations", 2nd Ed. Oxford Univ. Press (London and New York), 1958.

29. P. O. Löwdin, J. Mol. Spectroscopy, 14, 131 (1964).

30. P. D. Robinson, Proc. Phys. Soc. (London) 78, 537 (1961), considered a very interesting related problem. Using a delta function model for the electronic potential energy, he considered the diatomic hydrogen ion, H_2^+, as a hydrogen atom polarized by a proton. Robinson found that the resultant perturbation sequence did not converge for large values of the internuclear separation, R. This result is very reasonable since the series solution cannot decide whether to converge to the gerade or the ungerade state of the diatomic hydrogen ion. For small values of R, Robinson's series does converge to the gerade (ground) state. In a similar manner, in electron exchange problems, it may be necessary to use zeroth order functions with the proper symmetry in order to insure that the perturbation series will converge to the proper state of the system.

31. J. I. Musher, "Approximate Wave Functions for Molecules and Crystals", submitted to Revs. Mod. Phys. (October, 1965).

32. H. Primas, Istanbul International Summer School of Quantum Chemistry; 1964 (Academic Press, 1965).

33. A. A. Ovchinnikov and A. D. Sukhanov, Physics Doklady 9, 685 (1965), require that their series expansion (for the diatomic hydrogen ion, H_2^+) have certain desirable properties. The resulting consistency conditions are then used to determine the expansion parameter.

34. J. O. Hirschfelder, J. Chem. Phys. 39, 2099 (1963); R. A. Sack, J. Chem. Phys. 39, 3275 (1963); R. Yaris, J. Chem. Phys. 40, 1891 (1964).

35. P. D. Robinson, J. H. Huntington, and J. O. Hirschfelder, J. Chem. Phys. 42, 1879 (1965).

36. P. D. Robinson, University of Wisconsin, Mathematics Research Center, U. S. Army, Report #609 (December, 1965).

37. J. O. Hirschfelder, J. Chem. Phys. 00, 0000 (December 15, 1965).

This research was supported by the National Aeronautics Space Administration Grant NsG-275-62.